PROCEEDINGS, TOUGH Symposium 2003
Lawrence Berkeley National Laboratory, Berkeley, California, May 12–14, 2003

COMPARATIVE SIMULATIONS OF A LARGE-SCALE FIELD INFILTRATION EXPERIMENT

Wenbin Wang[1], Shlomo P. Neuman[2] , Tzung-mow Yao[3], and Peter J. Wierenga[3]

[1]Now with Water Management Consultants, 3025 N Campbell Ave, #281, Tucson, AZ 85719.
e-mail: wwang@watermc.com
[2]Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721
e-mail: neuman@hwr.arizona.edu
[3]Department of Soil, Water, and Environmental Science, University of Arizona, Tucson, AZ 85721
e-mail: myao@ag.arizona.edu. wierenga@ag.arizona.edu

ABSTRACT

TOUGH2 and iTOUGH2 are used to conduct forward and inverse simulations of a large-scale infiltration experiment at the Maricopa Agricultural Center (MAC) near Phoenix, Arizona. Three site representations are considered: uniform horizontal layers, layers composed of uniform segments, and layers having randomly varying properties. Due to a paucity of hydraulic parameter measurements at the MAC, these are inferred from soil compositional data using generic data bases and pedotransfer functions. Variogram analyses of these data support the laterally nonuniform and randomly variable site representations. To reproduce observed water contents it is necessary to modify the inferred hydraulic parameters through inverse modeling based on preliminary sensitivity and error analyses. Model discrimination criteria are used to rank the three calibrated site models. Layers composed of uniform segments are ranked highest due to their superior performance and relative simplicity. The choice representation is validated by using it to reproduce water contents during another infiltration experiment.

INTRODUCTION

Analyses of water flow in the vadose zone are often hampered by lack of adequate site data. Without such data, it is difficult to develop detailed predictive models of unsaturated flow in heterogeneous soils. Under what circumstances can relatively simple models, based on data that are relatively easy to obtain, provide reliable predictions of flow in the vadose zone? What levels of hydrogeological complexity and detail are required to reproduce observed groundwater flow at a vadose zone site? How reliably can flow at a site be reproduced by means of simple models, and how simple can such models be? These questions are acutely relevant to those charged with environmental safety assessment in common situations where time and resources are severely limited. To address them, we conduct comparative simulations of a large-scale field infiltration experiment at the Maricopa Agricultural Center (MAC) near Phoenix, Arizona, using site models of increasing complexity.

EXPERIMENTS AND SITE CHARACTERIZATION

Brief Description of Infiltration Experiments

Three large-scale infiltration experiments were conducted at the MAC (Young *et al.*, 1999) by applying water uniformly at a controlled rate to a 50 × 50 m^2 area using a drip irrigation system. The area was covered by a 60 × 60 m^2 liner to minimize evaporation. We consider the first and last of these three experiments. Experiment 1 lasted 93 days starting April 28 and ending July 30, 1997. Water was applied at an average rate of 1.85 *cm/day* to the field for 24 days, with a bromide tracer added for the first 15 days at a mean concentration of 31.6 *ppm*. The water application period was followed by a redistribution period of 69 days. Experiment 3 lasted more than 200 days; we consider the first 56 days starting April 24 and ending June 19, 2001. Water was applied at an average rate of 2.66 *cm/day* for 28 days and redistribution measured for the following 28 days. Monitoring took place across the site using a variety of devices among which we concentrate on neutron probe readings of water content taken in 9 boreholes down to a depth of 14 meters at 0.25 *m* intervals (see Figure 1).

Site characterization and alternative representations

The MAC site is situated in a basin filled with alluvial deposits ranging from hundreds to thousands of feet in depth. Down to a depth of 15 *m* the soils consist of sand, sandy loam, and loamy sand. Variograms of percent sand, silt, and clay show a vertical range of 2 *m*, which provides support for a layered representation of the soils; the same is suggested by a visual examination of soil profiles in boreholes and neutron count ratios, which correlate with soil compositional data (Wang, 2002). The neutron data reveal a perched water table at a depth of about 13 *m*. Hence it appears sensible to represent the upper 13 *m* of the vadose zone by 8–9 layers consisting of three distinct soil types (sandy loam, gravelly loam sand, and sand) as illustrated in Figure 1.

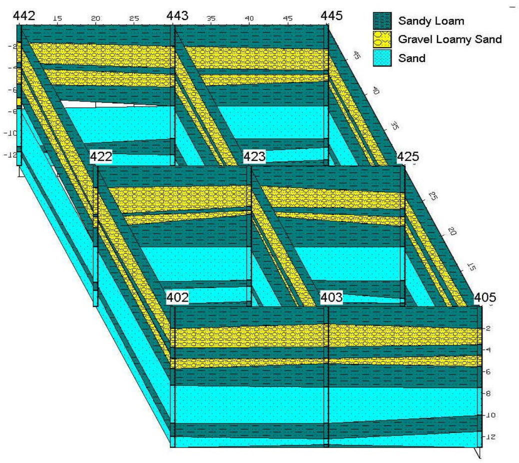

Figure 1. Local stratigraphy based on soil and neutron data (scale in meters).

Omni-directional variograms of percent sand, silt and clay in soil intervals of 30 *cm* down to 1.8 *m* exhibit horizontal ranges of 20 - 25 *m*. Hence another sensible site model is one in which the layers are segmented laterally into uniform zones measuring 20 – 25 *m*. Yet another plausible representation is one of layers having randomly varying soil compositions, with a structure represented by the corresponding variograms.

There is a paucity of hydraulic parameter measurements at the MAC. We therefore rely on pedotransfer functions (Bouma and van Lanen, 1997) to translate soil pedologic data into hydraulic parameters. We start by ascribing to each uniform layer hydraulic parameter values equal to mean values of a generic database for the corresponding soil class. Wang *et al.* (2003) considered three such databases (RAWLS, Rawls *et al.*, 1982; ROSETTA, Schaap and Leij, 1998; CARSEL, Carsel and Parrish, 1988, and Meyer *et al.*, 1997) and found that one of them (CARSEL) gave better reproductions of observed behavior at the MAC than the other two. We therefore ascribe CARSEL mean values to the layers.

Pedologic data associated with individual soil samples were translated by Wang (2002) into hydraulic parameters using the Rosetta neural network software of Schaap *et al.* (1998). Variograms of corresponding base-10 log saturated hydraulic conductivity (log K_s) estimates exhibited vertical ranges (down to a depth of 15.5 m) of 1 - 2 *m*, horizontal ranges (down to a depth of 1.8 *m*) of 20 - 36 *m* (mostly 20–25 *m*), sill of about 0.22 and nugget effect of about 0.07. This suggests a site representation in which log K_s varies randomly within each layer about the mean value assigned to it on the basis of CARSEL. Random variations about the mean are taken to be lognormal with a spherical variogram, vertical range of 2 m, horizontal range of 25 m, and the above sill and nugget values. A single random field is generated for each layer using GSLIB (Deutsch and Journel, 1998). Point values at the centers of grid blocks are assigned to the blocks. All other hydraulic parameters remain equal to their mean values.

FORWARD FLOW MODELING

Forward simulations are conducted using TOUGH2 (Pruess *et al.*, 1999) by assigning the above mean or random hydraulic parameters to each cell of a finite difference grid. Cells measure 2.0 *m* horizontally and 10 cm vertically, covering a N-S vertical section 13 *m* deep and 110 *m*. The section passes through boreholes 402, 422, and 442 (Figure 1). The bottom boundary is a water table treated as a zero pressure boundary. No flow is allowed to take place across the lateral boundaries. During the first 28 days of experiment 3, a constant nonzero flux is prescribed within the irrigated plot and zero flux measured in boreholes 402, 422, and 442 prior to the experiment are taken to represent initial conditions.

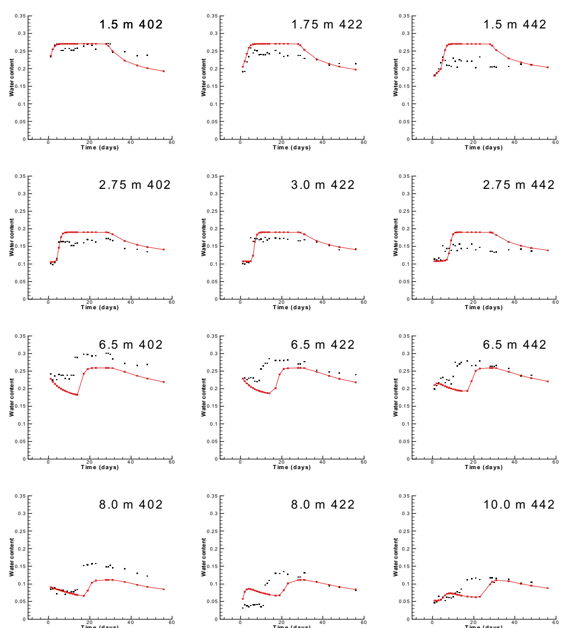

Figure 2. Forward simulation of water contents along N-S uniformly layered transect using mean hydraulic parameter values from CARSEL at various depths.

Figure 2 shows the results of forward simulation with uniform layers. The computed response (curves) captures in a very crude way the observed behavior (dots). The quality of the results varies with depth, and simulated wetting front arrival times are seen to lag by up to ten days behind those measured.

Segmented layers belong to one soil class and would therefore yield similar results. Forward simulation using a random realization of log saturated conductivities results in a slightly better reproduction of observed water contents. However, none of the forward simulations are satisfactory and there is an obvious need to calibrate the models against observed system behavior outside this plot at the top boundary. During the remainder of the experiment, the soil surface constitutes a no-flow boundary. Water contents

INVERSE FLOW MODELING

We use iTOUGH2 (Finsterle, 1999a-b) to calibrate our flow models against observed water contents at the MAC. The inverse code estimates hydraulic parameters by minimizing a weighted sum of squared water content and parameter residuals. Weights are set equal to inverse variances of water content measurement and parameter estimation errors, respectively. The standard deviation of a water content measurement error is taken to be 10% of the measured value. The variance of any mean CARSEL parameter is taken to be that of the same parameter in the database. In the case of randomly generated parameters, the sill and nugget are added to the variance. Preliminary analysis shows that model results are most sensitive to K_s and the van Genuchten (1980) parameters n and α, in that order (Table 1). Hence these are the parameters we estimate. The scaled sensitivity coefficients in the table are defined as:

$$\omega_j = \sigma_{p_j} \sum_{i=1}^{m} \left| \frac{\partial z_i}{\partial p_j} \frac{1}{\sigma_{z_i}} \right| \qquad (1)$$

where the standard deviation σ_{p_j} of j parameter adopted from Meyer's results (1997), m equals to the number of temporal data points, and the *a priori* standard deviations σ_{z_i} of the observation z_i are of 10% the measured values. Sensitivity analysis shows that K_S and n estimates exhibit strong negative correlation and that the calibration criterion is not very sensitive to K_s. We employ automatic parameter selection criteria provided by iTOUGH2 to help overcome these difficulties.

As sandy loam at depths 0 - 2 *m* has different bulk densities than do deeper sandy loam layers, we estimate its hydraulic properties separately. This yields a total of four materials: sandy loam in the top layer, sandy loam in deeper layers, gravel loamy sand and sand.

Table1. Results of Preliminary Sensitivity Analysis

Parameters	Scaled Sensitivity Coefficient.	Sens. of Calib. Criterion
log (k_s) LOAM1	1.71E+02	3.92E-01
log (k_s) GSAND	1.29E+02	8.53E-03
log (k_s) LOAM2	3.46E+01	4.58E-02
log (k_s) SAND	5.96E+01	1.34E-01
log (α) LOAM1	8.00E+00	4.51E-02
log (α) GSAND	4.02E+00	4.96E-01
log (α) LOAM2	3.30E+01	9.35E-01
log (α) SAND	2.13E-01	5.62E-01
log (λ) LOAM1	1.10E+02	4.33E+00
log (λ) GSAND	1.15E+02	6.85E+00
log (λ) LOAM2	3.72E+01	1.25E+00
log (λ) SAND	5.04E+01	2.07E-01

NOTE: LOAM1 represents sandy loam in the top layer, LOAM2 sandy loam in deeper layers, GSAND gravel loamy sand, and SAND sand. k_s represents permeability and λ=1-1/n, The standard deviation of hydraulic parameters from CARSEL database is adopted as scaling factor for sensitivity coefficient.

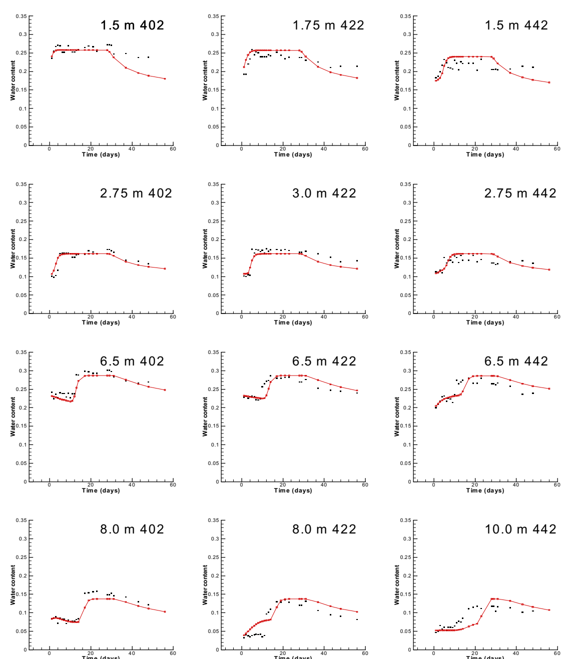

Figure 3. Simulation of water contents during infiltration experiment 3 along N-S transect using inverse estimates of saturated hydraulic conductivity and van Genuchten's α and n. Uniform soil layers.

Figure 3 depicts matches between simulated (curves) and measured (dots) water contents along the N-S cross-section by considering the soil to consist of horizontally uniform layers. The matches are seen to be much better than those obtained prior to inversion in Figure 2. While some simulation results fit the data well, other are systematically too low or too high. For example, water content in the top sandy loam layer in borehole 402 is systematically under-predicted, whereas in sandy loam and sand layers at depths 6 - 10 *m* in boreholes

Variogram analysis has shown that the dominant horizontal correlation scale of soil hydraulic parameters at the MAC is 20 - 25 *m*. We therefore subdivide the transect into 3 horizontal segments, one per borehole. This yields a total of 36 parameters for the transect.

Figure 4 compares simulated and observed water contents using inverse parameter estimates along the transect. The fit is seen to be good in all cases. A histogram of residuals (Figure 5) suggests that they are close to normal with a near-zero mean and small standard deviation. At a confidence level of 95%, only 14 out of the 300 residuals are identified as outliers.

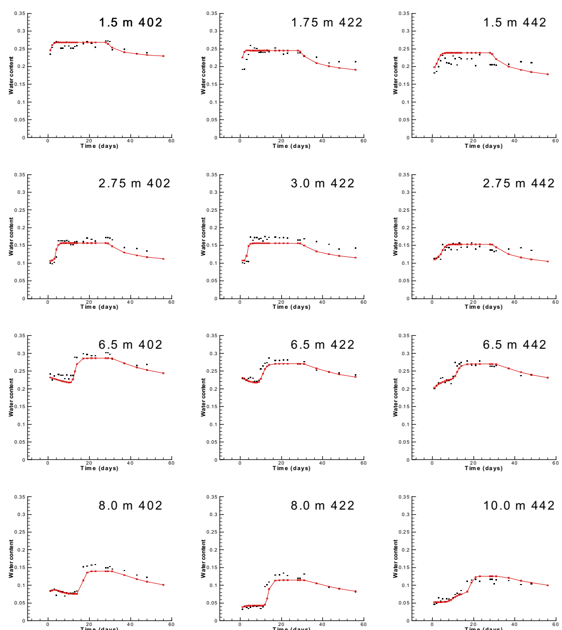

Figure 4. Simulation of water contents during infiltration experiment 3 along N-S transect using inverse estimates of saturated hydraulic conductivity and van Genuchten's α and n. Segmented layers.

Figure 6 compares simulated and observed water contents following inversion when the layers are considered to be randomly heterogeneous. In this case we consider each layer to have a spatially uniform mean, which we estimate using the above inverse procedure. We then superimpose on the estimated mean log hydraulic conductivity of each layer a random fluctuation, which is not affected by the inverse procedure. The result is almost the same as that for laterally segmented layers in Figure 5. This is so despite a large difference between the corresponding permeability estimates, as shown in Figure 7. There is little channeling due to a pronounced layering effect.

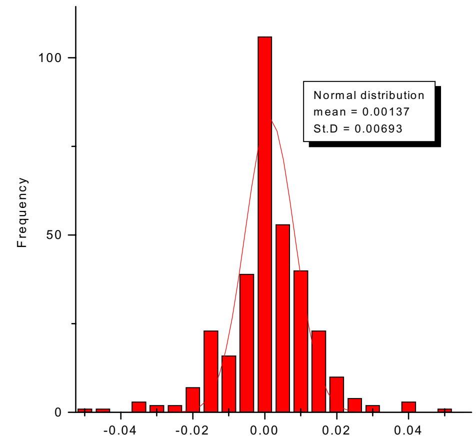

Figure 5. Histogram of differences between observed and simulated water contents along N-S transect following inversion. Segmented layers

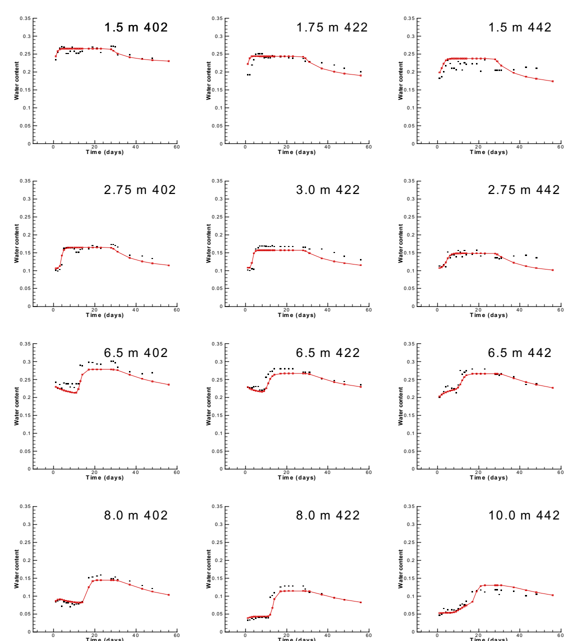

Figure 6. Simulation of water contents during infiltration experiment 3 along N-S transect using inverse estimates of saturated hydraulic conductivity and van Genuchten's α and n. Randomly heterogeneous layers

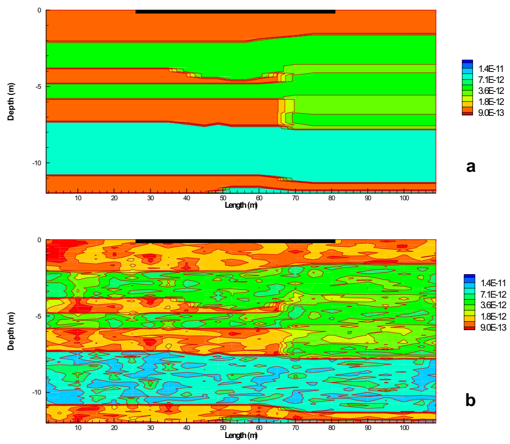

Figure 7. Inverse permeability estimates corresponding to a) segmented and b) randomly heterogeneous layers.

Table 2. Model Quality Criteria.

Model Quality Criteria	2-D N-S models		
	Uniform layers	Segmented layers	Randomly hetero. layers
D-optimality	5.83E-57	5.50E-179	8.69E-174
A-optimality	1.96E-03	1.11E-03	2.04E-03
E-optimality	8.49E-04	2.35E-04	1.11E-03
Sum of weighted squared residuals	8.20E+02	2.70E+02	2.631E+02
Log of the latter	6.11E+02	1.00E+03	9.65E+02
Akaike	-1.20E+03	-1.93E+03	-1.86E+03
Kashyap	-1.04E+03	-1.45E+03	-1.45E+3

Tables 2 compares the three sets of inverse results. The quality of model fit is compared on the basis of a D-optimality criterion (determinant of the covariance matrix of parameter estimation errors), A-optimality criterion (trace of this matrix), E-optimality criterion (largest absolute eigenvalue of the same matrix, Steinberg and Hunter, 1984), calibration criterion (weighted sum of squared residuals) and its logarithm. These model fit criteria can be used to compare the quality of different models having similar structure and number of parameters but not models having different structures or numbers of parameters. To validly compare the quality of all models in Tables 2, we employ formal model discrimination criteria due to Akaike (1974) and Kashyap (1982) as done previously by Carrera and Neuman (1986a-c). The smaller (or more negative) are these criteria, the better is the model. Both model discrimination criteria identify the uniformly layered model as being the worst. The segmented and randomly heterogeneous models perform equally well. Considering that the randomly heterogeneous model is much more complex than the segmented model, the latter is our obvious choice for purposes of predicting water content variations at the MAC. We suspect, however, that a randomly heterogeneous model may be a better choice for purposes of simulating solute transport at the site.

CONFIRMATION OF INVERSE MODELING RESULTS

Our choice site representation (segmented layers) and inverse hydraulic parameter estimates are based on data collected during infiltration experiment 3. We use them to simulate water contents at the MAC during experiment 1. The good matches between simulated and observed water contents in Figure 8 constitute a confirmation of the calibrated model.

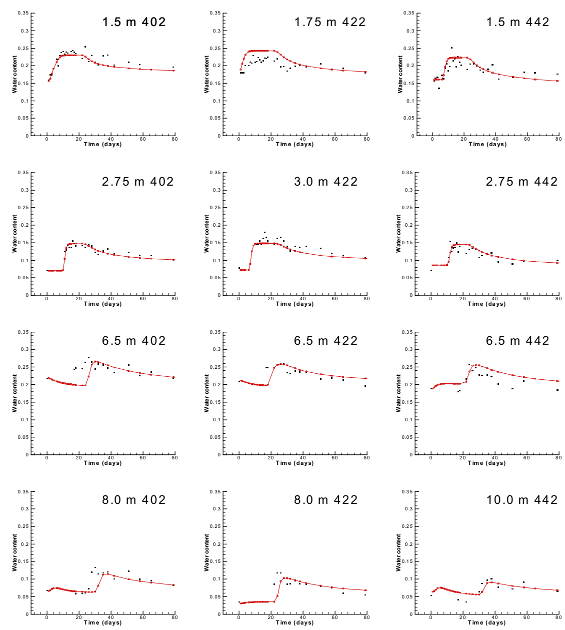

Figure 8. Simulation of water contents during experiment 1 along N-S transect (boreholes 402, 422, 442) using segmented layer model and hydraulic parameters obtained from experiment 3.

CONCLUSIONS

Our paper leads to the following conclusions:

1. Hydraulic parameter estimates, obtained from a generic database or by means of pedotransfer functions, are unable to reproduce water contents observed during a large-scale infiltration experiment at the MAC without calibrating the flow model against such observations.
2. Three site representations are considered: uniform horizontal layers, layers composed of uniform segments, and layers having randomly varying properties. Model discrimination criteria identify the segmented model as best among the three for purposes of predicting water content variations at the MAC due to its superior performance and relative simplicity. We suspect, however, that a randomly heterogeneous model may be a better choice for purposes of simulating solute transport at the site.
3. The latter model, calibrated against water content data observed during infiltration experiment 3, reproduces with fidelity water content data observed during infiltration experiment 1.

ACKNOWLEDGMENT

This work was supported by the US Nuclear Regulatory Commission under contract number NRC-04-97-056.

REFERENCES

Akaike, H., New look at statistical model identification, *IEEE Trans. Autom. Control*. AC19(6), 716–723, 1974.

Bouma, J., and J. A. van Lanen, Transfer functions and threshold values: From soil characteristics to land qualities, In *Quantified land evaluation*, edited by K.J. Beek et al. pp. 106-110, International Institute Aerospace Surv. Earth Sci. ITC publ., 6, Enschede, Netherlands, 1987.

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 1, Maximum likelihood incorporating prior information, *Water Resour. Res*., 22(2), 199–210, 1986a.

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 2, Uniqueness, stability, and solution algorithms, *Water Resour. Res.*, 22(2), 211–227, 1986b.

Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 3, Application to synthetic and field data, *Water Resour. Res.*, 22(2), 228–242, 1986c.

Carsel, R. F., and R. S. Parrish., Developing joint probability distributions of soil water retention characteristics, *Wat . Resour. Res.*, 24(5), 755-770, 1988.

Deutsch, C. V., and A.G. Journel, *GSLIB – Geostatistical Software Library and User’s Guide*, Second Ed., Applied Geostatistical Series, Oxford University Press, New York Oxford, 1997.

Finsterle, S., *iTOUGH2 User's Guide*, Report LBNL-40040, Lawrence Berkeley National Laboratory, Berkeley, Calif., 1999a.

Finsterle, S., *iTOUGH2 Sample Problems*, Report LBNL-40042, Lawrence Berkeley National Laboratory, Berkeley, Calif., 1999b.

Kashyap, R. L., Optimal choice of AR and MA parts in auto regressive moving average models, *IEEE Trans. Pattern Anal. Mach. Intell.* 4(2), 99–104, 1982.

Meyer, P. D., M. L. Rockhold, and G. W. Gee, *Uncertainty Analyses of Infiltration and Subsurface Flow and Transport for SDMP Sites*, Report NUREG/CR-6565, U.S. Nuclear Regulatory Commission, Washington, DC, 1997.

Pruess, K., C. Oldendurg, and G. Moridis, *TOUGH2 user's guide, Version 2*, Report LBNL-43134, Lawrence Berkeley National Laboratory, Berkeley, 1999.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton, Estimation of soil water properties, *Trans. ASAE* 25(5), 1316-1320, 1982.

Schaap, M.G., and F.J. Leij, Database Related Accuracy and Uncertainty of Pedotransfer Functions,. *Soil Science*, 163(10), 765-779, 1998.

Schaap, M.G., F.J. Leij, and M.T. van Genuchten, Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity, *Soil Sci Soc Am J.* 62(4), 847-855, 1998.

Steinberg, D. M., and W. G. Hunter, Experimental design: Review and comment, *Technometrics*, 28, 71-97, 1984.

van Genuchten, M. T, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci Soc Am J*., 44(5), 892–898, 1980.

Wang, W., *Uncertainty Analysis of Groundwater Flow and Solute Transport in Unsaturated-Saturated Soil: Maricopa Case*, Ph.D. Dissertation, University of Arizona, 2002.

Wang, W, S. P. Neuman, T. Yao, and P. J. Wierenga, Simulation of Large-Scale Field Infiltration Experiments Using a Hierarchy of Models Based on Public, Generic and Site Data, *J. of Vadoze Zone*, 2003 (in Press).

Young, M.H., P. J. Wierenga, A. W. Warrick, L. L. Hofmann, S. A. Musil, M. Yao, C. J. Mai, Z. Zou, and B. R. Scanlon, *Results of field studies at the Maricopa Environmental Monitoring Site, Arizona*, Report NUREG/CR-5694. U.S. Nuclear Regulatory Commission, Washington, D.C, 1999.